K-Means Clustering of People with COVID-19

July 3, 2020

DoYeong, Jeon
Software Engineering Laboratory
Soongsil University

TABLE OF CONTENTS

1. Source Code

1.1. Code for Creating Database

❑ **CreatingDB Class**

```
class CreatingDB:
    """
    Class for creating random database
    """
    num_people = 0  # number of people to create
    base_date = None  # the base date of data

    def __init__(self, num_people, base_date):
        self.num_people = num_people
        self.base_date = base_date

    def generate_incurred_date(self):
        """
        function to create random incurred date
        :return:
            incurred_date: string, the day of infection or contact
            elapsed_days: int, the difference between base date and incurred
date
        """
        elapsed_days = random.randint(0, 14)  # the valid day period is 0~14
        # extracting the incurred day using periods and base date
        incurred_date = (self.base_date - timedelta(days=elapsed_days)). \
            strftime("%Y %m %d")
        return incurred_date, elapsed_days

    def generate_address_list(self):
        """
        function to get one address randomly from the adress list
        :return: the randomly generated address list
        """
        with open('./Address_Part.txt', 'r', encoding='utf-8') as add_file:
            # add_file = add_file.encoding
            address_list = add_file.readlines()

            random_address_list = []  # list to store addresses

            # extract addresses as many as the number of recipients
            for _ in range(1, self.num_people + 1):
                random_address_list.append(random.choice(address_list))

        return random_address_list

    def generate_csv_data(self):
        """
        function to create .csv file with randomly generated records
        :return: None
        """
```

```
        num_healthy = round(self.num_people / 3)  # 1/3 is healthy
        num_contacted = round(self.num_people / 3)  # 1/3 is contacted
        # 1/3 is confirmed
        num_confirmed = self.num_people - num_healthy - num_contacted

        id_list = list(range(1, self.num_people + 1))  # ID as many as people
        random.shuffle(id_list)  # shuffle list

        # age records as many as people
        age_list = list(random.randint(1, 100)
                        for _ in range(1, self.num_people + 1))
        # address records as many as people
        address_list = self.generate_address_list()

        severity_list = []  # severity records as many as people
        incurred_date_list = []  # incurred date list including 'None'(healthy)
        status_list = []  # status(Healthy, Contacted, and Confirmed) list

        # Entire people num = healthy + contacted + confirmed
        # Repeat as many healthy people
        for _ in range(num_healthy):
            # severity_list.append(0)
            status_list.append('Healthy')
            incurred_date_list.append('None')

        # Repeat as many contacted people
        for count in range(num_contacted):
            date, days = self.generate_incurred_date()
            status_list.append('Contacted')
            #     severity_list.append(round(self.compute_severity('contacted',
days), 2))
            incurred_date_list.append(date)

        # Repeat as many confirmed people
        for _ in range(num_confirmed):
            date, days = self.generate_incurred_date()
            status_list.append('Confirmed')
            #     severity_list.append(round(self.compute_severity('confirmed',
days), 2))
            incurred_date_list.append(date)

        # converting as pandas DataFrame data type to save .csv
        df = pd.DataFrame({
            "ID": id_list,
            "Age": age_list,
            "Address": address_list,
            "Covid Status": status_list,
            # "Severity": severity_list,
            "Incurred Date": incurred_date_list,
        })
        df = df.sort_values(['ID'], ascending=[True])
        df.reset_index(drop=True, inplace=True)

        # saving as .csv file
```

```
        df.to_csv("corona_data.csv", mode='w', encoding='utf-8-sig')
```

1.2. Code for Clustering

❑ **ClusteringPeople Class**

```
class ClusteringPeople:
    df_corona = None
    cluster_result_dic = {}
    num_healthy = 0
    healthy_id_list = []
    num_contacted = 0
    contacted_id_list = []
    num_confirmed = 0
    confirmed_id_list = []

    def __init__(self, file_path):
        self.load_data(file_path)
        self.preprocess()
        self.compute_people_number_of_type()

    def load_data(self, file_path):
        """
        method to load .csv file
        :param file_path: string, the path of file
        :return:
        """
        self.df_corona = pd.read_csv(file_path)

    def compute_people_number_of_type(self):
        status_series = self.df_corona["Covid Status"]
        for idx in range(len(status_series)):
            if status_series[idx] == 'Contacted':
                self.num_contacted += 1
                self.contacted_id_list.append(idx+1)
            elif status_series[idx] == 'Confirmed':
                self.num_confirmed += 1
                self.confirmed_id_list.append(idx+1)
            else:
                self.num_healthy += 1
                self.healthy_id_list.append(idx+1)

    def compute_average_severity(self, id_list):
        sum_of_severity = 0
        for id in id_list:
            sum_of_severity += self.df_corona["Severity"][id-1]
        return sum_of_severity / len(id_list)


    def display_load_data(self):
        print(f"Total number of People: {len(self.df_corona)}")
        print(f"{'ID':<4}{'Age':<4}{'Covid
```

```
Status':<13}{'Severity':<9}{'Address':<10}")
        for i in range(len(self.df_corona)):
            print(f"{self.df_corona['ID'][i]:<4}"
                  f"{self.df_corona['Age'][i]:<4}"
                  f"{self.df_corona['Covid Status'][i]:<13}"
                  f"{round(self.df_corona['Severity'][i], 3):<9}"
                  f"{self.df_corona['Address'][i].split()[0]:<10}"
                  )
        print()  # float 1 line
        print(f"Number of healthy people: {self.num_healthy}")
        print(f"Number of contacted people: {self.num_contacted}")
        print(f"Number of confirmed people: {self.num_confirmed}")
        print(f"Average Severity of contacted people: "
              f"{round(self.compute_average_severity(self.contacted_id_list),
2)}")
        print(f"Average Severity of confirmed people: "
              f"{round(self.compute_average_severity(self.confirmed_id_list),
2)}")
        print()  # float 1 line

    def preprocess(self):
        """
        method to preprocess the data for distance function
        :return: None
        """
        col_num = len(self.df_corona)  # the number of rows from loaded data
        today = datetime.now().date()  # date of today, YEAR-MONTH-DAY

        # selecting specific column to compute 'severity'
        incur_date_col = self.df_corona['Incurred Date']
        status = self.df_corona['Covid Status']

        severity_list = []  # list for storing severity result

        for i in range(col_num):
            severity = 0  # default is healthy, 0.
            if status[i] == 'Contacted':  # contacted person?
                # formula for contacted person:
                #   x = 1 - ((today's date) - (infected date)) * 0.05)
                elapsed_days = (today - parse(incur_date_col[i]).date()).days
                severity = (1 - (elapsed_days * 0.05)) * 0.5

            elif status[i] == 'Confirmed':  # confirmed person?
                # formula for confirmed person:
                #   x = (1 - ((today's date) - (infected date)) * 0.05)) / 2
                elapsed_days = (today - parse(incur_date_col[i]).date()).days
                severity = 1 - (elapsed_days * 0.05)

            severity_list.append(severity)  # add the value to the list
        self.df_corona["Severity"] = severity_list

    def cluster(self):
        sse_list = []  # list for storing SSE(Sum of squares errors)
        silhouette_score_list = []  # list for storing silhouette scores
```

```
        for i in range(2, 10):  # number of clusters 2 to 9
            # load the k-means model
            km = cluster.KMeans(
                n_clusters=i,  # the number of cluster
                init='k-means++',  # how to initial cluster centers
                max_iter=300,  # maximum number of iterations
                algorithm='auto'  # three choices: auto, full, and elkan.
            )

            # changing the shape of data
            severity_list = self.df_corona["Severity"].values.tolist()
            severity_list = np.array(severity_list)

            # cluster
            cluster_predicted_list  =  km.fit_predict(severity_list.reshape(-1,
1))

            # storing SSE value to get the optimal number of cluster
            sse_list.append(km.inertia_)

            # storing silhouette score to get optimal number of cluster

silhouette_score_list.append(silhouette_score(severity_list.reshape(-1,   1),
cluster_predicted_list))

            cluster_list = [j for j in range(i)]  # cluster list
            # display the reuslt of cluster
            self.print_result_of_cluster(cluster_list, cluster_predicted_list)

            # store the prediction result
            self.cluster_result_dic[i] = cluster_predicted_list

    def draw_elbow_method(self, sse_list):
        """
        method to draw elbow graph using SSE(Sum of Squares Error)
        :param sse_list: list of SSE
        :return: None
        """
        plt.plot(range(2, 10), sse_list, marker='o')
        plt.xlabel("The Number of Cluster")
        plt.ylabel("SSE")
        plt.show()

    def print_result_of_cluster(self, cluster_list, cluster_predicted_list):
        severity_list = self.df_corona["Severity"].values.tolist()
        id_list = self.df_corona["ID"].values.tolist()

        cluster_predicted_list = cluster_predicted_list.tolist()
        people_num_of_each_cluster_list = []
        avg_severity_of_each_cluster_list = []

        print(f"Number of Clusters: {len(cluster_list)}")
        for cluster_idx in cluster_list:  # 1 cluster
```

```
            num_people = cluster_predicted_list.count(cluster_idx)
            id_severity_tuple_list = []
            sum_of_severities = 0
            for person_idx in range(len(cluster_predicted_list)):
                if cluster_idx == cluster_predicted_list[person_idx]:
                    sum_of_severities += severity_list[person_idx]
                    id_severity_tuple_list.append((person_idx+1,
round(severity_list[person_idx], 2)))

            people_num_of_each_cluster_list.append(num_people)

            print(f"\tCluster {cluster_idx}:")
            print(f"\t\tNumber of People: {num_people}")
            print(f"\t\t\t{'ID':<4}{'Severity Value'}")
            for person_in_cluster in id_severity_tuple_list:

print(f"\t\t\t{person_in_cluster[0]:<4}{person_in_cluster[1]}")

            print(f"\t\tAverage of severities: {round(sum_of_severities /
len(id_severity_tuple_list), 2)}")
            avg_severity_of_each_cluster_list.append(round(sum_of_severities /
len(id_severity_tuple_list), 2))

            print()  # float 1 line
        self.display_table(people_num_of_each_cluster_list,
                           avg_severity_of_each_cluster_list)
        print()  # float 1 line

    def display_table(self,
                      people_of_cluster_list,
                      avg_severity_of_cluster_list):
        print(f"\t{'-'*42}")
        print(f"\t{'Cluster ID':^11}|{' # of People':^11}| {'Avg of
Severity':^15}")
        cluster_id = 0
        for people_num, avg in zip(people_of_cluster_list,
avg_severity_of_cluster_list):
            print(f"\t{cluster_id:^11}| {people_num:>5} | {avg:<10}")
            cluster_id += 1
        print(f"\t{'-'*11}|{'-'*12}|{'-'*17}")
        print(f"\t{'Total':^11}| {sum(people_of_cluster_list):^11}|")
        print(f"\t{'-'*42}")

    def draw_silhouette(self):
        """
        method to draw graph using silhouette scores
        :return: None
        """
        pass

    def draw_graph(self):
        """
        method to draw clustering result
        :return: None
```

```
        """
        pass
```

❑ **main**

```
if __name__ == '__main__':
    # CODE FOR CLUSTERING
    file_path = './corona_data.csv'

    cp = ClusteringPeople(file_path)
    cp.preprocess()
    cp.draw_graph()
    cp.cluster()
```

2. Result of Clustering

2.1. Loaded Dataset

❑ **Top 25 lines**

```
Total number of People: 100
ID  Age Covid Status Severity Address
1   72  Contacted    0.15     충청남도
2   50  Healthy      0.0      경기도
3   49  Contacted    0.225    경상북도
4   45  Contacted    0.25     전라남도
5   45  Contacted    0.35     전라남도
6   66  Confirmed    0.6      부산광역시
7   86  Healthy      0.0      전라남도
8   43  Healthy      0.0      서울특별시
9   63  Healthy      0.0      경기도
10  81  Confirmed    0.4      광주광역시
11  2   Contacted    0.2      경상북도
12  69  Healthy      0.0      전라북도
13  66  Healthy      0.0      전라북도
14  37  Contacted    0.45     울산광역시
15  97  Healthy      0.0      경상북도
16  98  Healthy      0.0      경상북도
17  56  Confirmed    0.7      전라북도
18  26  Contacted    0.325    경상남도
19  90  Confirmed    0.95     전라북도
20  21  Confirmed    0.55     전라북도
21  26  Healthy      0.0      경상북도
22  17  Confirmed    0.5      제주특별자치도
23  55  Healthy      0.0      부산광역시
24  74  Healthy      0.0      경상북도
25  91  Contacted    0.3      경상북도
```

❑ **Last 25 lines and Statistics**

```
76  72  Contacted    0.4      경상북도
77  67  Confirmed    0.4      서울특별시
78  16  Healthy      0.0      경기도
79  62  Contacted    0.475    전라북도
80  24  Healthy      0.0      경상북도
81  10  Confirmed    0.6      대구광역시
82  72  Confirmed    0.9      인천광역시
83  70  Contacted    0.225    경기도
84  30  Confirmed    0.65     경기도
85  37  Healthy      0.0      경상북도
86  23  Contacted    0.175    전라북도
87  13  Confirmed    0.6      경상남도
88  34  Confirmed    0.3      대구광역시
89  19  Confirmed    0.3      충청북도
90  12  Healthy      0.0      전라남도
91  88  Healthy      0.0      대구광역시
92  80  Healthy      0.0      충청북도
93  13  Healthy      0.0      서울특별시
94  46  Confirmed    0.7      서울특별시
95  49  Contacted    0.35     서울특별시
96  15  Confirmed    0.3      경기도
97  37  Confirmed    0.7      경기도
98  40  Healthy      0.0      경상남도
99  65  Confirmed    0.6      충청북도
100 45  Confirmed    0.7      충청남도

Number of healthy people: 33
Number of contacted people: 33
Number of confirmed people: 34
Average Severity of contacted people: 0.3
Average Severity of confirmed people: 0.62
```

2.2. K-Means

- **Number of Clusters: 2**
  - Cluster 0, Top 25 lines

```
Number of Clusters: 2
    Cluster 0:
        Number of People: 70
            ID  Severity Value
            1   0.15
            2   0.0
            3   0.22
            4   0.25
            5   0.35
            7   0.0
            8   0.0
            9   0.0
            10  0.4
            11  0.2
            12  0.0
            13  0.0
            15  0.0
            16  0.0
            18  0.32
            21  0.0
            23  0.0
            24  0.0
            25  0.3
```

- Cluster 0, last 10 lines and average

```
85  0.0
86  0.17
88  0.3
89  0.3
90  0.0
91  0.0
92  0.0
93  0.0
95  0.35
96  0.3
98  0.0
Average of severities: 0.16
```

- Cluster 1

```
Cluster 1:
    Number of People: 30
        ID  Severity Value
        6   0.6
        14  0.45
        17  0.7
        19  0.95
        20  0.55
        22  0.5
        28  0.95
        29  0.5
        31  0.45
        36  0.95
        40  1.0
        44  0.47
        46  0.65
        47  0.8
        54  0.95
        56  0.47
        59  0.9
        62  0.5
        63  0.75
        66  0.55
        68  0.5
        79  0.47
        81  0.6
        82  0.9
        84  0.65
        87  0.6
        94  0.7
        97  0.7
        99  0.6
        100 0.7
    Average of severities: 0.67
```

- Summary Table

```
-----------------------------------------
Cluster ID | # of People| Avg of Severity
     0     |      70    |      0.16
     1     |      30    |      0.67
-----------|------------|-----------------
   Total   |     100    |
-----------------------------------------
```

- **Number of Clusters: 3**
  - Cluster 0, top 10 lines

```
Number of Clusters: 3
    Cluster 0:
        Number of People: 44
            ID  Severity Value
            3   0.22
            4   0.25
            5   0.35
            10  0.4
            11  0.2
            14  0.45
            18  0.32
            20  0.55
            22  0.5
            25  0.3
```

  - Cluster 0, last 10 lines

```
            70  0.4
            74  0.35
            75  0.35
            76  0.4
            77  0.4
            79  0.47
            83  0.22
            88  0.3
            89  0.3
            95  0.35
            96  0.3
        Average of severities: 0.36
```

  - Cluster 1

```
Cluster 1:
    Number of People: 19
        ID  Severity Value
        6   0.6
        17  0.7
        19  0.95
        28  0.95
        36  0.95
        40  1.0
        46  0.65
        47  0.8
        54  0.95
        59  0.9
        63  0.75
        81  0.6
        82  0.9
        84  0.65
        87  0.6
        94  0.7
        97  0.7
        99  0.6
        100 0.7
    Average of severities: 0.77
```

- Cluster 2

```
Cluster 2:
    Number of People: 37
        ID  Severity Value
        1   0.15
        2   0.0
        7   0.0
        8   0.0
        9   0.0
        12  0.0
        13  0.0
        15  0.0
        16  0.0
        21  0.0
        23  0.0
        24  0.0
        26  0.0
        27  0.0
        35  0.0
        39  0.0
        41  0.0
        42  0.15
        43  0.0
        48  0.0
        49  0.0
        50  0.0
        51  0.0
        58  0.0
        60  0.15
        71  0.0
        72  0.0
        73  0.0
        78  0.0
        80  0.0
        85  0.0
        86  0.17
        90  0.0
        91  0.0
        92  0.0
        93  0.0
        98  0.0
    Average of severities: 0.02
```

- Summary Table

Cluster ID	# of People	Avg of Severity
0	44	0.36
1	19	0.77
2	37	0.02
Total	100	

- **Number of Clusters: 4**
  - Cluster 0

```
Number of Clusters: 4
    Cluster 0:
        Number of People: 36
            ID  Severity Value
            1   0.15
            2   0.0
            7   0.0
            8   0.0
            9   0.0
            12  0.0
            13  0.0
            15  0.0
            16  0.0
            21  0.0
            23  0.0
            24  0.0
            26  0.0
            27  0.0
            35  0.0
            39  0.0
            41  0.0
            42  0.15
            43  0.0
            48  0.0
            49  0.0
            50  0.0
            51  0.0
            58  0.0
            60  0.15
            71  0.0
            72  0.0
            73  0.0
            78  0.0
            80  0.0
            85  0.0
            90  0.0
            91  0.0
            92  0.0
            93  0.0
            98  0.0
        Average of severities: 0.01
```

- Cluster 1

```
    Cluster 1:
        Number of People: 9
            ID  Severity Value
            19  0.95
            28  0.95
            36  0.95
            40  1.0
            47  0.8
            54  0.95
            59  0.9
            63  0.75
            82  0.9
        Average of severities: 0.91
```

- Cluster 2

```
Cluster 2:
    Number of People: 34
        ID  Severity Value
        3   0.22
        4   0.25
        5   0.35
        10  0.4
        11  0.2
        18  0.32
        25  0.3
        30  0.3
        32  0.35
        33  0.4
        34  0.2
        37  0.32
        38  0.35
        45  0.2
        52  0.4
        53  0.35
        55  0.25
        57  0.32
        61  0.3
        64  0.3
        65  0.22
        67  0.3
        69  0.25
        70  0.4
        74  0.35
        75  0.35
        76  0.4
        77  0.4
        83  0.22
        86  0.17
        88  0.3
        89  0.3
        95  0.35
        96  0.3
    Average of severities: 0.31
```

- Cluster 3

```
Cluster 3:
    Number of People: 21
        ID  Severity Value
        6   0.6
        14  0.45
        17  0.7
        20  0.55
        22  0.5
        29  0.5
        31  0.45
        44  0.47
        46  0.65
        56  0.47
        62  0.5
        66  0.55
        68  0.5
        79  0.47
        81  0.6
        84  0.65
        87  0.6
        94  0.7
        97  0.7
        99  0.6
        100 0.7
    Average of severities: 0.57
```

- Summary Table

Cluster ID	# of People	Avg of Severity
0	36	0.01
1	9	0.91
2	34	0.31
3	21	0.57
Total	100	

- **Number of Clusters: 5**
  - Cluster 0, 1

```
Number of Clusters: 5
    Cluster 0:
        Number of People: 13
            ID  Severity Value
            6   0.6
            17  0.7
            20  0.55
            46  0.65
            63  0.75
            66  0.55
            81  0.6
            84  0.65
            87  0.6
            94  0.7
            97  0.7
            99  0.6
            100 0.7
        Average of severities: 0.64

    Cluster 1:
        Number of People: 21
            ID  Severity Value
            1   0.15
            3   0.22
            4   0.25
            11  0.2
            25  0.3
            30  0.3
            34  0.2
            42  0.15
            45  0.2
            55  0.25
            60  0.15
            61  0.3
            64  0.3
            65  0.22
            67  0.3
            69  0.25
            83  0.22
            86  0.17
            88  0.3
            89  0.3
            96  0.3
        Average of severities: 0.24
```

- Cluster 2, 3

```
Cluster 2:
    Number of People: 33
        ID  Severity Value
        2   0.0
        7   0.0
        8   0.0
        9   0.0
        12  0.0
        13  0.0
        15  0.0
        16  0.0
        21  0.0
        23  0.0
        24  0.0
        26  0.0
        27  0.0
        35  0.0
        39  0.0
        41  0.0
        43  0.0
        48  0.0
        49  0.0
        50  0.0
        51  0.0
        58  0.0
        71  0.0
        72  0.0
        73  0.0
        78  0.0
        80  0.0
        85  0.0
        90  0.0
        91  0.0
        92  0.0
        93  0.0
        98  0.0
    Average of severities: 0.0

Cluster 3:
    Number of People: 8
        ID  Severity Value
        19  0.95
        28  0.95
        36  0.95
        40  1.0
        47  0.8
        54  0.95
        59  0.9
        82  0.9
    Average of severities: 0.93
```

- Cluster 4

```
Cluster 4:
    Number of People: 25
        ID  Severity Value
        5   0.35
        10  0.4
        14  0.45
        18  0.32
        22  0.5
        29  0.5
        31  0.45
        32  0.35
        33  0.4
        37  0.32
        38  0.35
        44  0.47
        52  0.4
        53  0.35
        56  0.47
        57  0.32
        62  0.5
        68  0.5
        70  0.4
        74  0.35
        75  0.35
        76  0.4
        77  0.4
        79  0.47
        95  0.35
    Average of severities: 0.41
```

- Summary Table

Cluster ID	# of People	Avg of Severity
0	13	0.64
1	21	0.24
2	33	0.0
3	8	0.93
4	25	0.41
Total	100	

- **Number of Clusters: 6**
  - Cluster 0

```
Number of Clusters: 6
    Cluster 0:
        Number of People: 33
            ID  Severity Value
            2   0.0
            7   0.0
            8   0.0
            9   0.0
            12  0.0
            13  0.0
            15  0.0
            16  0.0
            21  0.0
            23  0.0
            24  0.0
            26  0.0
            27  0.0
            35  0.0
            39  0.0
            41  0.0
            43  0.0
            48  0.0
            49  0.0
            50  0.0
            51  0.0
            58  0.0
            71  0.0
            72  0.0
            73  0.0
            78  0.0
            80  0.0
            85  0.0
            90  0.0
            91  0.0
            92  0.0
            93  0.0
            98  0.0
        Average of severities: 0.0
```

- Cluster 1, 2

```
Cluster 1:
    Number of People: 11
        ID  Severity Value
        6   0.6
        17  0.7
        46  0.65
        63  0.75
        81  0.6
        84  0.65
        87  0.6
        94  0.7
        97  0.7
        99  0.6
        100 0.7
    Average of severities: 0.66

Cluster 2:
    Number of People: 24
        ID  Severity Value
        5   0.35
        10  0.4
        18  0.32
        25  0.3
        30  0.3
        32  0.35
        33  0.4
        37  0.32
        38  0.35
        52  0.4
        53  0.35
        57  0.32
        61  0.3
        64  0.3
        67  0.3
        70  0.4
        74  0.35
        75  0.35
        76  0.4
        77  0.4
        88  0.3
        89  0.3
        95  0.35
        96  0.3
    Average of severities: 0.34
```

- Cluster 3

```
Cluster 3:
    Number of People: 8
        ID  Severity Value
        19  0.95
        28  0.95
        36  0.95
        40  1.0
        47  0.8
        54  0.95
        59  0.9
        82  0.9
    Average of severities: 0.93
```

- Cluster 4, 5

```
Cluster 4:
    Number of People: 13
        ID  Severity Value
        1   0.15
        3   0.22
        4   0.25
        11  0.2
        34  0.2
        42  0.15
        45  0.2
        55  0.25
        60  0.15
        65  0.22
        69  0.25
        83  0.22
        86  0.17
    Average of severities: 0.2

Cluster 5:
    Number of People: 11
        ID  Severity Value
        14  0.45
        20  0.55
        22  0.5
        29  0.5
        31  0.45
        44  0.47
        56  0.47
        62  0.5
        66  0.55
        68  0.5
        79  0.47
    Average of severities: 0.49
```

- Summary Table

Cluster ID	# of People	Avg of Severity
0	33	0.0
1	11	0.66
2	24	0.34
3	8	0.93
4	13	0.2
5	11	0.49
Total	100	

- **Number of Clusters: 7**
  - Cluster 0

```
Number of Clusters: 7
    Cluster 0:
        Number of People: 8
            ID  Severity Value
            6   0.6
            20  0.55
            46  0.65
            66  0.55
            81  0.6
            84  0.65
            87  0.6
            99  0.6
        Average of severities: 0.6
```

- Cluster 1

```
    Cluster 1:
        Number of People: 33
            ID  Severity Value
            2   0.0
            7   0.0
            8   0.0
            9   0.0
            12  0.0
            13  0.0
            15  0.0
            16  0.0
            21  0.0
            23  0.0
            24  0.0
            26  0.0
            27  0.0
            35  0.0
            39  0.0
            41  0.0
            43  0.0
            48  0.0
            49  0.0
            50  0.0
            51  0.0
            58  0.0
            71  0.0
            72  0.0
            73  0.0
            78  0.0
            80  0.0
            85  0.0
            90  0.0
            91  0.0
            92  0.0
            93  0.0
            98  0.0
        Average of severities: 0.0
```

- Cluster 2, 3

```
Cluster 2:
    Number of People: 18
        ID  Severity Value
        5   0.35
        18  0.32
        25  0.3
        30  0.3
        32  0.35
        37  0.32
        38  0.35
        53  0.35
        57  0.32
        61  0.3
        64  0.3
        67  0.3
        74  0.35
        75  0.35
        88  0.3
        89  0.3
        95  0.35
        96  0.3
    Average of severities: 0.32

Cluster 3:
    Number of People: 7
        ID  Severity Value
        19  0.95
        28  0.95
        36  0.95
        40  1.0
        54  0.95
        59  0.9
        82  0.9
    Average of severities: 0.94
```

- Cluster 4

```
Cluster 4:
    Number of People: 13
        ID  Severity Value
        1   0.15
        3   0.22
        4   0.25
        11  0.2
        34  0.2
        42  0.15
        45  0.2
        55  0.25
        60  0.15
        65  0.22
        69  0.25
        83  0.22
        86  0.17
    Average of severities: 0.2
```

- Cluster 5, 6

```
Cluster 5:
    Number of People: 15
        ID  Severity Value
        10  0.4
        14  0.45
        22  0.5
        29  0.5
        31  0.45
        33  0.4
        44  0.47
        52  0.4
        56  0.47
        62  0.5
        68  0.5
        70  0.4
        76  0.4
        77  0.4
        79  0.47
    Average of severities: 0.45

Cluster 6:
    Number of People: 6
        ID  Severity Value
        17  0.7
        47  0.8
        63  0.75
        94  0.7
        97  0.7
        100 0.7
    Average of severities: 0.73
```

- Cluster 5, 6

```
Cluster 5:
    Number of People: 12
    People list with Severity Values:
        ID  Severity Value
        5   0.7
        25  0.6
        37  0.65
        61  0.6
        67  0.6
        75  0.7
    Average of severities: 0.65
Cluster 6:
    Number of People: 4
    People list with Severity Values:
        ID  Severity Value
        14  0.9
        56  0.95
    Average of severities: 0.94
```

- Summary Table

Cluster ID	# of People	Avg of Severity
0	8	0.6
1	33	0.0
2	18	0.32
3	7	0.94
4	13	0.2
5	15	0.45
6	6	0.73
Total	100	

- **Number of Clusters: 8**
  - Cluster 0, 1

```
Number of Clusters: 8
    Cluster 0:
        Number of People: 14
            ID  Severity Value
            4   0.25
            18  0.32
            25  0.3
            30  0.3
            37  0.32
            55  0.25
            57  0.32
            61  0.3
            64  0.3
            67  0.3
            69  0.25
            88  0.3
            89  0.3
            96  0.3
        Average of severities: 0.29

    Cluster 1:
        Number of People: 6
            ID  Severity Value
            17  0.7
            47  0.8
            63  0.75
            94  0.7
            97  0.7
            100 0.7
        Average of severities: 0.73
```

  - Cluster 2

```
Cluster 2:
    Number of People: 33
        ID  Severity Value
        2   0.0
        7   0.0
        8   0.0
        9   0.0
        12  0.0
        13  0.0
        15  0.0
        16  0.0
        21  0.0
        23  0.0
        24  0.0
        26  0.0
        27  0.0
        35  0.0
        39  0.0
        41  0.0
        43  0.0
        48  0.0
        49  0.0
        50  0.0
        51  0.0
        58  0.0
        71  0.0
        72  0.0
        73  0.0
        78  0.0
        80  0.0
        85  0.0
        90  0.0
        91  0.0
        92  0.0
        93  0.0
        98  0.0
    Average of severities: 0.0
```

- Cluster 3, 4, 5

```
Cluster 3:
    Number of People: 9
        ID  Severity Value
        14  0.45
        22  0.5
        29  0.5
        31  0.45
        44  0.47
        56  0.47
        62  0.5
        68  0.5
        79  0.47
    Average of severities: 0.48

Cluster 4:
    Number of People: 7
        ID  Severity Value
        19  0.95
        28  0.95
        36  0.95
        40  1.0
        54  0.95
        59  0.9
        82  0.9
    Average of severities: 0.94

Cluster 5:
    Number of People: 10
        ID  Severity Value
        1   0.15
        3   0.22
        11  0.2
        34  0.2
        42  0.15
        45  0.2
        60  0.15
        65  0.22
        83  0.22
        86  0.17
    Average of severities: 0.19
```

- Cluster 6, 7

```
Cluster 6:
    Number of People: 8
        ID  Severity Value
        6   0.6
        20  0.55
        46  0.65
        66  0.55
        81  0.6
        84  0.65
        87  0.6
        99  0.6
    Average of severities: 0.6

Cluster 7:
    Number of People: 13
        ID  Severity Value
        5   0.35
        10  0.4
        32  0.35
        33  0.4
        38  0.35
        52  0.4
        53  0.35
        70  0.4
        74  0.35
        75  0.35
        76  0.4
        77  0.4
        95  0.35
    Average of severities: 0.37
```

- Summary Table

Cluster ID	# of People	Avg of Severity
0	14	0.29
1	6	0.73
2	33	0.0
3	9	0.48
4	7	0.94
5	10	0.19
6	8	0.6
7	13	0.37
Total	100	

- **Number of Clusters: 9**
  - Cluster 0

```
Number of Clusters: 9
    Cluster 0:
        Number of People: 33
            ID  Severity Value
            2   0.0
            7   0.0
            8   0.0
            9   0.0
            12  0.0
            13  0.0
            15  0.0
            16  0.0
            21  0.0
            23  0.0
            24  0.0
            26  0.0
            27  0.0
            35  0.0
            39  0.0
            41  0.0
            43  0.0
            48  0.0
            49  0.0
            50  0.0
            51  0.0
            58  0.0
            71  0.0
            72  0.0
            73  0.0
            78  0.0
            80  0.0
            85  0.0
            90  0.0
            91  0.0
            92  0.0
            93  0.0
            98  0.0
        Average of severities: 0.0
```

- Cluster 1, 2, 3, 4

```
Cluster 1:
    Number of People: 8
        ID  Severity Value
        10  0.4
        14  0.45
        31  0.45
        33  0.4
        52  0.4
        70  0.4
        76  0.4
        77  0.4
    Average of severities: 0.41

Cluster 2:
    Number of People: 7
        ID  Severity Value
        19  0.95
        28  0.95
        36  0.95
        40  1.0
        54  0.95
        59  0.9
        82  0.9
    Average of severities: 0.94

Cluster 3:
    Number of People: 6
        ID  Severity Value
        3   0.22
        4   0.25
        55  0.25
        65  0.22
        69  0.25
        83  0.22
    Average of severities: 0.24

Cluster 4:
    Number of People: 6
        ID  Severity Value
        6   0.6
        46  0.65
        81  0.6
        84  0.65
        87  0.6
        99  0.6
    Average of severities: 0.62
```

- Cluster 5, 6

```
Cluster 5:
    Number of People: 9
        ID  Severity Value
        20  0.55
        22  0.5
        29  0.5
        44  0.47
        56  0.47
        62  0.5
        66  0.55
        68  0.5
        79  0.47
    Average of severities: 0.5

Cluster 6:
    Number of People: 18
        ID  Severity Value
        5   0.35
        18  0.32
        25  0.3
        30  0.3
        32  0.35
        37  0.32
        38  0.35
        53  0.35
        57  0.32
        61  0.3
        64  0.3
        67  0.3
        74  0.35
        75  0.35
        88  0.3
        89  0.3
        95  0.35
        96  0.3
    Average of severities: 0.32
```

- Cluster 7, 8

```
Cluster 7:
    Number of People: 7
        ID  Severity Value
        1   0.15
        11  0.2
        34  0.2
        42  0.15
        45  0.2
        60  0.15
        86  0.17
    Average of severities: 0.17

Cluster 8:
    Number of People: 6
        ID  Severity Value
        17  0.7
        47  0.8
        63  0.75
        94  0.7
        97  0.7
        100 0.7
    Average of severities: 0.73
```

- Summary Table

Cluster ID	# of People	Avg of Severity
0	33	0.0
1	8	0.41
2	7	0.94
3	6	0.24
4	6	0.62
5	9	0.5
6	18	0.32
7	7	0.17
8	6	0.73
Total	100	